

## Genetic Analysis of Complex Traits

Jing Hua Zhao, MRC Epidemiology Unit, Cambridge, UK

Large quantity of genetic data is currently available to investigate their association with complex traits such as common human diseases and other quantitative measurements with contributions of many genes of small effects. There have been extensive international collaborations and vigorous developments in appropriate statistical methods and computational tools.

The tutorial begins with an outline of genetic dissection of complex traits with focus on design and analysis in genetic association studies particularly genome-wide association studies (GWAS) which involve large number of individuals and GeneChip data containing ~1,000,000 or more single nucleotide polymorphisms, the most abundant genetic variants in human genome. Topics range from quality controls and descriptive analysis to assessing genotype-phenotype relationship and inference of pathways. Genotype imputation, meta-analysis and graphical presentation will be illustrated with case studies involving population-based and family-based samples and use of data in publicly available projects, such as HapMap, 1000 genomes and dbGaP. They will enable the participants become familiar with the computational and statistical problems, popular specialized software and relevant packages in R, e.g., *genetics*, *haplo.stats*, *gap*, *kinship*, as with packages designed for GWAS such as *GenABEL*, *snpMatrix*, *SNPassoc*, and *NCBI2R*.

The tutorial is based on considerable theoretical and practical work in statistical genetics and genetic epidemiology, especially design and analysis for GWAS in several large epidemiological cohorts and contributions to a variety of consortia. It incorporates materials from a series of presentations including two previous useR! tutorials and will be appropriate for researchers with basic knowledge in statistics and computing and wish to get involved with genetic data analysis in humans while generating interest to researchers in plant and animal sciences. It will also be useful to professionals and researchers actively engaged in analysis of genetic data and/or development of computational tools in environments such as R.

## References

Altshuler D, MJ Daly, ES Lander. Genetic mapping in human disease. *Science* 322:881-888, 2008

Hindorff LA, P Sethupathy, HA Junkins, EM Ramosa, JP. Mehtac, FS Collins, TA Manolio. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proceedings of National Academy of Science* 106(23):9362-7, 2009

Lin S, H Zhao (Eds). Handbook on Analyzing Human Genetic Data: Computational Approaches and Software. Springer, 2010

Zhao JH, Q Tan. Integrated analysis of genetic data with R. *Human Genomics*. 2:258-265, 2006

# Schedule

**(Tuesday 20/7/2010 9:30am – 13:00pm)**

---

9:30am	–	10:15am	Overview
10:15am	–	11:00am	Association Analysis
11:00am	–	11:30am	Break
11:30am	–	12:15pm	Meta-Analysis
12:15pm	–	13:00pm	Other issues

---

The tutorial will be lecture only and laptop is not required. It has a web site:

[www.mrc-epid.cam.ac.uk/~jinghua.zhao/notes/useR2010](http://www.mrc-epid.cam.ac.uk/~jinghua.zhao/notes/useR2010)